高被引论文的在线使用与分享研究*
——基于 ALMs 的实证分析

匡登辉
(南开大学图书馆　天津 300071)

摘要:【目的】以 SCI 高被引论文为实证分析对象，通过分析论文的引用频次与在线使用、分享的相关性及表现，验证 Altmetrics 指标在揭示优质文献方面的有效性。【方法】通过匹配 DOI，将高被引论文的引文数据与 ALMs 组合，对论文的引用频次与在线使用、分享数据进行 Spearman 相关性检验，并对 PLOS 系列期刊的在线使用与分享情况逐一分析。【结果】研究表明：引用次数与论文的在线分享 Figshare 呈弱正相关($r = 0.081$, $p = 0.01$)；与 CiteULike 标引量呈中度相关($r = 0.252$, $p = 0.01$)；与 Mendeley 使用量的相关性最大($r = 0.376$, $p = 0.01$)。部分期刊的 Mendeley 使用量在一定程度上揭示了高影响力文献。【局限】只针对特定学科 PLOS 系列期刊，因此数据的全面性存在一定的欠缺，得出的结论能否推而广之，有待于进一步研究。【结论】论文的高被引与其对应的在线使用与分享的相关度不高，显示出在线使用(Mendeley，CiteULike)与分享(Figshare)的数据在揭示高影响力文献方面存在一定的不足。

关键词: 高被引论文　在线使用　Mendeley　CiteULike　分享　Figshare
分类号: G250

1 引 言

引文是滞后的文献计量指标。例如在物理学领域，发表后的文章一般需要 5 年才获得被引用、被参考的信息[1]，这样就推迟了科研评价的进程。而且，“重要的论文可能被更频繁地引用”，只是基于一种理想化的假设，大家对这种文献计量即采用引文分析用于评价保持慎重态度[2-3]。虽然引文分析在当下仍不失为一项重要研究评估因素，但是单靠引文已不能有效衡量学术交流在各种新层面的拓展：如读者在线评论研究成果及与科研人员在社交网络互动交流等。Altmetrics 的广泛应用，正逐渐凸显传统计量的局限。

Altmetrics 可追踪任何形式发表的研究成果，超越了传统的文章的影响因子和引文分析的评价，填补了一般的引用评价指标时间延迟的不足，丰富完善了科研评价体系，全面评价了学术成果的影响力，包括其学术影响力和社会影响力两方面。

对科研人员来说，Altmetrics 将大数据转换成有用的信息：研究者可以从中了解针对科研所做的推广与努力，发现其他相关的科研同行，全球哪些地区、哪些研究者对相关领域感兴趣，对相关的研究及成果进行交流、互动；下次的研究成果应该发表在哪里，才能获得更广泛的传播和认可；以及潜在的科研合作者的学术表现如何。因此，Altmetrics 的兴起与科学交流的社交化、网络化密切相关。

对图书馆员来说，图书馆被行政主管部门要求提供关于教师著作的计量工作，日益增加，现在教师申报各类各级人才、申请奖励等，也需要提供重要创新性论文被知名索引数据库的收录、引用证明文件，而机构典藏(Institutional Repository, IR)与科研信息管理系统(Current Research Information System, CRIS)，作为存取、展示和传播科研信息的工具，在计量评价活

通讯作者：匡登辉，ORCID：0000-0001-7771-6830，E-mail：dhkuang@nankai.edu.cn。
*本文系 2016 年南开大学亚洲研究中心资助项目“高品质科技期刊的 Altmetrics 分析”(项目编号：AS 1606)的研究成果之一。

动中扮演重要角色。许多查新馆员参与、讨论各种科学计量评价, 包括传统的收录、引用, 也包括使用 Altmetrics, 并且在科研机构对 Altmetrics 进行解说和传播, Altmetrics 为图书馆提供了一种新的方式来表达其存在价值[4]。

作为 Web 2.0 环境下基于社交网络学术交互行为的影响力计量方法, Altmetrics 实时高效追踪学术成果的传播, 精确计量各种用户不同的在线交互, 测度学术成果的多元化影响力。论文的收录引用情况固然在一定程度上可以表征高影响力文献, Altmetrics 作为社交网络学术交互行为的影响力计量方法, 其相关指标能否与传统的引文分析指标(影响因子、被引频次等)一样, 揭示高影响力文献? 高被引文献是否具有高在线使用与分享指标? 或者反过来, 在线使用与分享指标高的文献, 是否与传统意义上高被引文献一致? 这是本文主要研究的三个问题。

本文利用 Altmetrics 数据集, 组配高被引论文数据, 选择其中的在线使用与分享指标, 分析其在整体上和各期刊层面的表现, 评估其在科学评价中的有效性, 同时也对高被引论文的在线使用与分享情况进行研究, 考察其能否替代被引频次而成为高影响力文献的表征指标, 即其能否揭示高影响力文献。

2　相关研究

2.1　Altmetrics 相关研究

Altmetrics 研究始于 2008 年, Taraborelli 在影响因子作为主要评价指标被质疑后, 呼吁建立一种基于社会软件的分布式的科学评价[5]; 2009 年, Neylon 和 Wu 以 PLOS 和 Faculty of 1000 为例, 从计量数据的来源和专家评议的激励机制两个角度论述论文层面的计量(Article-Level Metrics, ALMs)评价科学影响力方案的可行性[6]; 2010 年 7 月, Priem 和 Hemminger 提出基于社交网络的科学计量学 2.0(Scientometrics 2.0)概念, 并总结科学计量学 2.0 研究的各种类型的学术社交网络数据资源, 分析基于微博、社会化书签等 Web 2.0 应用提供的数据, 多角度地反映论文的影响力[7]。同年 10 月, Priem 等在 Altmetrics.org 发表“Altmetrics: A Manifesto”, 正式提出 Altmetrics[8]。2013 年, Galligan 等对 Altmetrics 的内涵进行阐释, 认为它是评价学术内容影响力的新方法, 并以学术内容在社交网络(如 Twitter)、文献在线管理工具(如 Mendeley, CiteULike)、分享平台(Figshare)中传播的广度作为评价的依据[9]。Altmetrics 已有一些具体实践应用, 典型的包括: Altmetric.com, Impact Story, Plum Analytics 和 PLOS 的 ALMs[10]。ALMs 是多维指标体系, 包含引文数据、论文使用(在线浏览和下载)和 Altmetrics(社交数据: 书签和传播活动、媒体和博客报道、社区讨论和评级)[11]。

2.2　论文的在线使用与分享研究

论文的在线使用与分享, 属于“社交数据: 书签和传播活动”, 是学术引用的“前奏”, 可以即时测量出学术论文的影响力, 从社会影响方面映射科学成果, 是 Altmetrics 的重要指标。本文从论文的在线使用与分享入手, 对这些 Altmetrics 指标在科学评价中的有效性进行评估。

国内外对于 Altmetrics 指标中论文的在线使用与分享的研究已经展开。其中, Ebrahimy 等考察了论文的在线使用与分享(CiteULike, Mendeley 和 Figshare)与 Scopus, Web of Science, PubMed Central 和 CrossRef 论文被引频次之间的相关性, 结果表明, CiteULike 和 Mendeley 与调查的所有引文系统的被引次数具有正相关, Figshare 与被引次数呈显著负相关[12]。Bornmann, Haustein 等, Li 等分别就不同的论文样本, 重点研究 Twitter, Mendeley 和 CiteULike 与传统论文的被引次数(Scopus, Web of Science)之间的相关性[13-15]; Bornmann 就 F1000 推荐的 PLOS 期刊论文(1 082 篇), 研究这些具有专家推荐标签的论文在 Facebook, Twitter, Mendeley, Figshare 的不同表现, 从而验证 Altmetrics 指标在评价科研成果方面的价值[16]。刘春丽等, 由庆斌等, 宋丽萍等, 刘晓娟等, 金玮等, Mohammadi 等分别就 Mendeley 指标在论文评价方面的价值进行实证研究[17-22]。Shrivastava 等就 Top 100 高被引物理论文的引用次数与 Mendeley 读者数之间的相关性进行研究[23]。这些研究主要就部分 Altmetrics 指标验证了与引文分析的相关性与一致性, 未就这些指标对高影响力文献的揭示作用进行研究, 而本文基于 ALMs 数据, 通过 DOI 映射关系, 匹配 ALMs 与高被引论文数据, 构建 SCI 高被引论文的 Altmetrics 数据集, 研究高被引论文在线使用与分享(Mendeley, CiteULike 和 Figshare)的情况, 从社会影响方面来评价科研产出, 考察论文的在线使用与分享数据在揭示高

影响力文献方面的有效性。

Figshare 是一种在线数字仓储，所有文件根据 DOI 进行分类存储。研究者可以保存和分享研究成果，包括数据、数据集、图像和影音等[24-25]，上传和下载内容均免费。CiteULike 允许用户保存和分享学术论文的引用，是基于社会书签的在线服务，可以促进研究者之间分享科研资料。Mendeley 是免费的跨平台文献管理软件，也是在线的学术社交网络平台，可对文献进行添加与组织、阅读与标注、与全球同行协作、备份、同步与移动、云存储与发现等。

在线文献管理软件整合了社会书签和文献管理功能。因此，本文将 CiteULike 标注量和 Mendeley 使用量的数据统一为在线使用，将 Figshare 分享量的数据作为分享数据。

3 研究框架与方法

3.1 数据来源与处理

(1) 数据来源与说明

在 Web of Knowledge 网站，选择 Web of Science 核心合集数据库，基本检索界面，设定论文筛选条件为 2003 年–2015 年出版的 PLOS 期刊论文，根据检出结果按照引用次数降序排列，导出引用次数≥100 的"全记录"数据到 EndNote，共 1 939 条，然后将数据导入 Excel，仅保留"DOI"和"引用次数"，建立高被引论文数据集，数据采集日期为 2015 年 2 月 3 日。

笔者从 Figshare 平台下载 ALMs[26]，知识共享协议 CC-BY 4.0。

ALMs 数据集(alm_stats_2015-01-10.csv)包含 PLOS 出版的所有论文(142 170 条记录，截至 2015-01-10)。此数据集包括丰富的论文级计量数据，主要有论文基本信息：论文题名、期刊名、出版时间、DOI；使用统计(数据来自在 PLOS、PubMed 网站，以 COUNTER 3 记载的论文的浏览、下载量)；来自 CrossRef、PubMed 和 Scopus 学术引文库的论文被引数据；社会分享(Facebook，Twitter，Figshare)、学术性标注(Mendeley 和 CiteULike)、论文评论(PLOS Comments)、非学术性引用情况(在 Nature Blogs 指向单篇论文的博客提及)、文章报道和策划(Article Coverage 和 Article Coverage Curated)、学术产出作者标识符(ORCID)、维基百科(Wikipedia)、F1000 因子(F1000)等[27]。

(2) 数据处理

由于 ALMs 中存在其与 DOI 的映射关系，笔者利用 DOI 映射，构建 SCI 高被引论文的 Altmetrics 数据集。通过匹配和数据清洗，只有 DOI(10.1371/journal.pmed.0050038)无匹配项，所以共 1 938 条有效数据。

根据 DOI 的不同，将 *PLOS Computational Biology* (PCBI)作为期刊 A，*PLOS Clinical Trials* (PCTR)作为期刊 B，*PLOS Genetics* (PGEN)作为期刊 C，*PLOS Neglected Tropical Diseases* (PNTD)作为期刊 D，*PLOS ONE* (PONE)作为期刊 E，*PLOS Pathogens* (PPAT)作为期刊 F，*PLOS Biology* (PBIO)作为期刊 G，*PLOS Medicine* (PMED)作为期刊 H。这样，本文分析的期刊共 8 种，由于在数据集中，*PLOS Clinical Trials* 仅 1 条数据，数据量太小，故舍弃，只针对其余 7 种期刊进行研究。

3.2 研究方法与应用软件

(1) 研究方法

本文主要研究期刊高被引论文的在线使用(Mendeley 和 CiteULike)与分享(Figshare)三项 Altmetrics 指标与论文引用量之间的相关性，探讨传统科研评价和学术的社会影响力之间的联系，检验论文的在线使用与分享能否揭示高影响力文献。

1 938 条数据涉及的论文集中在 2003 年–2014 年。对三项 Altmetrics 指标和 SCI 引用指标进行描述性统计；通过 K-S 单样本检验(定量)和 Q-Q 图(定性)，检查数据是否符合正态分布；为检测学术成果科学影响和社会影响的关系，计算 SCI 引用次数和三项 Altmetrics 指标之间的 Spearman 相关性系数；最后，从各期刊层面，考察 PLOS 系列期刊论文在线使用与分享的变化。

(2) 应用软件

使用的统计分析软件为 IBM SPSS Statistics 22.0。SPSS 22.0 界面非常友好，包含"Statistics Base"、"Statistics Regression"和"Advanced Statistics"等功能模块，提供了专业的"查看"、"计算"和"预测"功能。本文中的统计图表和结果均由 SPSS 输出。

4 研究过程

4.1 SPSS 描述性统计分析

由于选择的 SCI 引用是引用次数超过 100 次，因此 SCI 引用数最小值是 100，最大值是 2 170。在所有

的文献数据中, Mendeley 使用量、CiteULike 标注量和 Figshare 分享量的最小值均为 0, 如表 1 所示:

表 1　指标描述性统计

Item	Minimum	Maximum	Mean	Std. Deviation
Mendeley	0	13 015	133.29	318.769
Figshare	0	167	14.80	12.968
CiteULike	0	389	5.09	12.020
SCI	100	2 170	180.66	145.784
N=1938				

其中, 0.62%的论文未在 Mendeley 上使用, 24.87%的论文未在 CiteULike 使用, 1.08%的论文没有在 Figshare 分享, 如图 1 所示。即在本研究构建的高被引论文数据集中, Mendeley 使用量非零数据最多。

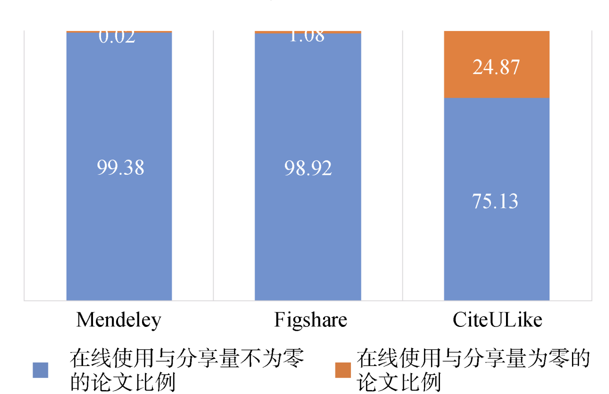


图 1　高被引论文的在线使用与分享统计

4.2　数据的正态分布检验

为选择合适的相关性检验方法, 需要对本研究中的 SCI 数据和论文的在线使用与分享数据进行正态性分布检验。主要选用定量分析的单样本 K-S 检验和 Q-Q 图定性分析来检验数据的正态性。

表 2 显示在线使用与分享计量指标、SCI 引用的双侧检验(P<0.05), 拒绝在线使用与分享指标、SCI 引用正态分布的假设, 因此该数据不符合正态分布。

表 2　单样本 K-S 检验

		Mendeley	Figshare	CiteULike	SCI
Normal Parameters[a, b]	Mean	133.29	14.80	5.09	180.66
	Std. Deviation	318.769	12.968	12.020	145.784
Most Extreme Differences	Absolute	.338	.173	.336	.290
	Positive	.268	.173	.252	.226
	Negative	–.338	–.151	–.336	–.290
Test Statistic		.338	.173	.336	.290
Asymp. Sig. (2-tailed)		.000[c]	.000[c]	.000[c]	.000[c]

(注: a. Test distribution is Normal (N=1938); b. Calculated from data; c. Lilliefors Significance Correction.)

SPSS 的基本统计分析功能里可以通过观察数据的 Q-Q 图判断数据是否服从正态分布。若该组数据服从正态分布, 则图中的点应该靠近图中直线, 而且该直线的斜率为标准差, 截距为均值。

以 Mendeley 为例, 图 2 显示曲线与直线相交, 并未接近重合, 因此该数据不符合正态分布, 所以不能使用 Pearson 方法检验相关性。同样的分析结果, 适用于 Figshare, CiteULike 和 SCI 的数据检验。

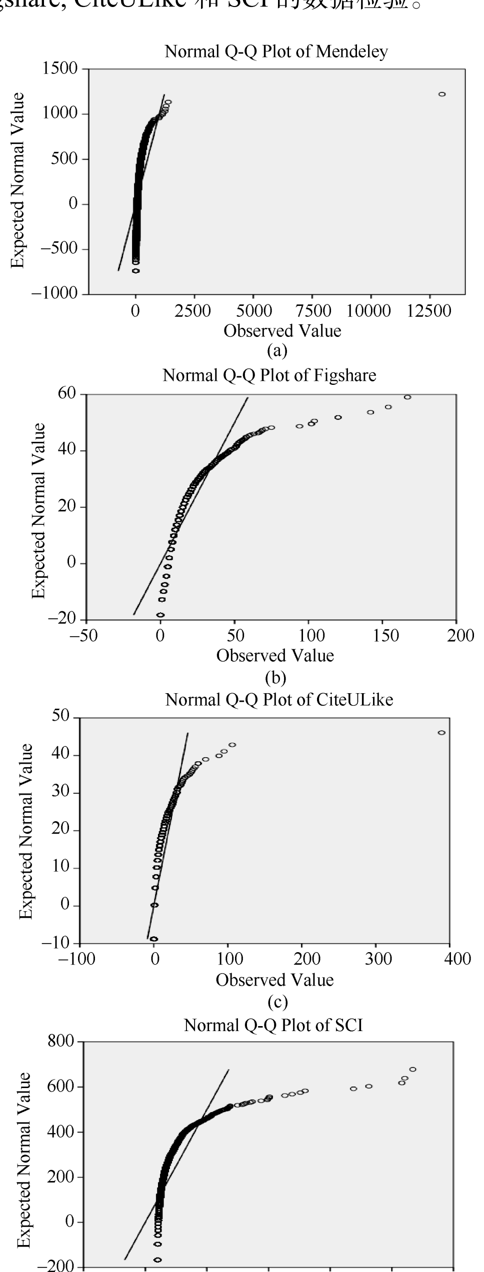


图 2　各指标的正态 Q-Q 图

4.3 相关性检验

笔者选择Spearman相关系数进行SCI引用和在线使用与分享计量指标之间的相关性检验，结果如表3所示：

表3 Spearman相关性检验

			Mendeley	Figshare	CiteULike	SCI
Spearman's rho N=1938	Mendeley	Correlation Coefficient	1.000	.061**	.596**	.376**
		Sig. (2-tailed)	.	.007	.000	.000
	Figshare	Correlation Coefficient	.061**	1.000	.063**	.081**
		Sig. (2-tailed)	.007	.	.006	.000
	CiteULike	Correlation Coefficient	.596**	.063**	1.000	.252**
		Sig. (2-tailed)	.000	.006	.	.000
	SCI	Correlation Coefficient	.376**	.081**	.252**	1.000
		Sig. (2-tailed)	.000	.000	.000	.

(**: Correlation is significant at the 0.01 level (2-tailed).)

SCI引用量和在线使用数与分享数存在正的弱相关，SCI引用与分享数(Figshare)的相关性很小($r = 0.081$, $p = 0.01$), SCI引用与CiteULike标注量的相关性居中($r = 0.252$, $p = 0.01$), Mendeley使用量与SCI引用之间相关性系数比较大($r = 0.376$, $p = 0.01$)。

矩阵散点图能够可视化地展示每对变量之间相关性。如图3所示，Mendeley在线使用和SCI引用具有较高的相关性。随着Mendeley在线使用的增加，SCI引用的次数也显著增加。也就是说，Mendeley在线使用极有可能导致未来论文引用次数的增加。从Spearman相关性系数计算和矩阵散点图分析，可以看出高被引论文在线使用与分享的数据用作科学影响力评价结果的一致性。

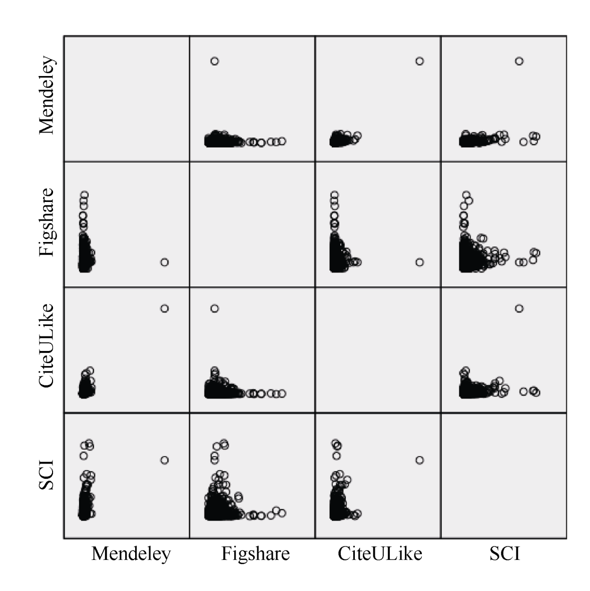


图3 矩阵散点图

4.4 PLOS系列期刊论文在线使用与分享的变化

上述关于论文在线使用与分享分析以及相关性检验是基于高被引数据和ALMs的组合数据集，即是基于PLOS期刊进行整体的分析研究，下面对高被引论文所属的期刊在线使用与分享情况进行分析。

从图4宏观上可以清晰地看出，Figshare和CiteULike的数值位于曲线底层，说明两平台上积累的用户数据量太小或者可能用户的活跃度太差。根据图4高被引论文所属的期刊在线使用与分享曲线的谷峰变化，PLOS期刊基本上可以分为两类，Mendeley峰值超过SCI峰值的期刊群(A, C, E, H), Mendeley峰值低于SCI峰值的期刊群(D, F, G)。对于前者，表明论文的在线使用量超过SCI引用次数，未来SCI的引用可能会进一步增加；对于后者，表明论文的在线使用量低于SCI引用次数，论文的在线使用并未成功揭示高影响力文献。

4.5 高被引论文的在线使用与分享分析

SCI引用次数前10的文献的在线使用与分享情况如表4所示。从被引次数前10的论文所属期刊统计，高被引论文分属于除A, B之外的6种期刊，其中期刊C(2篇)，期刊D(1篇)，期刊E(2篇)，期刊F(1篇)，期刊G(2篇)，期刊H(2篇)。

以Mendeley为例，在线使用前10的文献的SCI被引情况如表5所示。从Mendeley在线使用前10的论文所属期刊统计，论文分属于除A, B之外的6种期刊，其中期刊C(2篇)，期刊D(1篇)，期刊E(2篇)，期刊F(1篇)，期刊G(2篇)，期刊H(2篇)。从论文的期刊分布来看，表4和表5的结果高度一致。

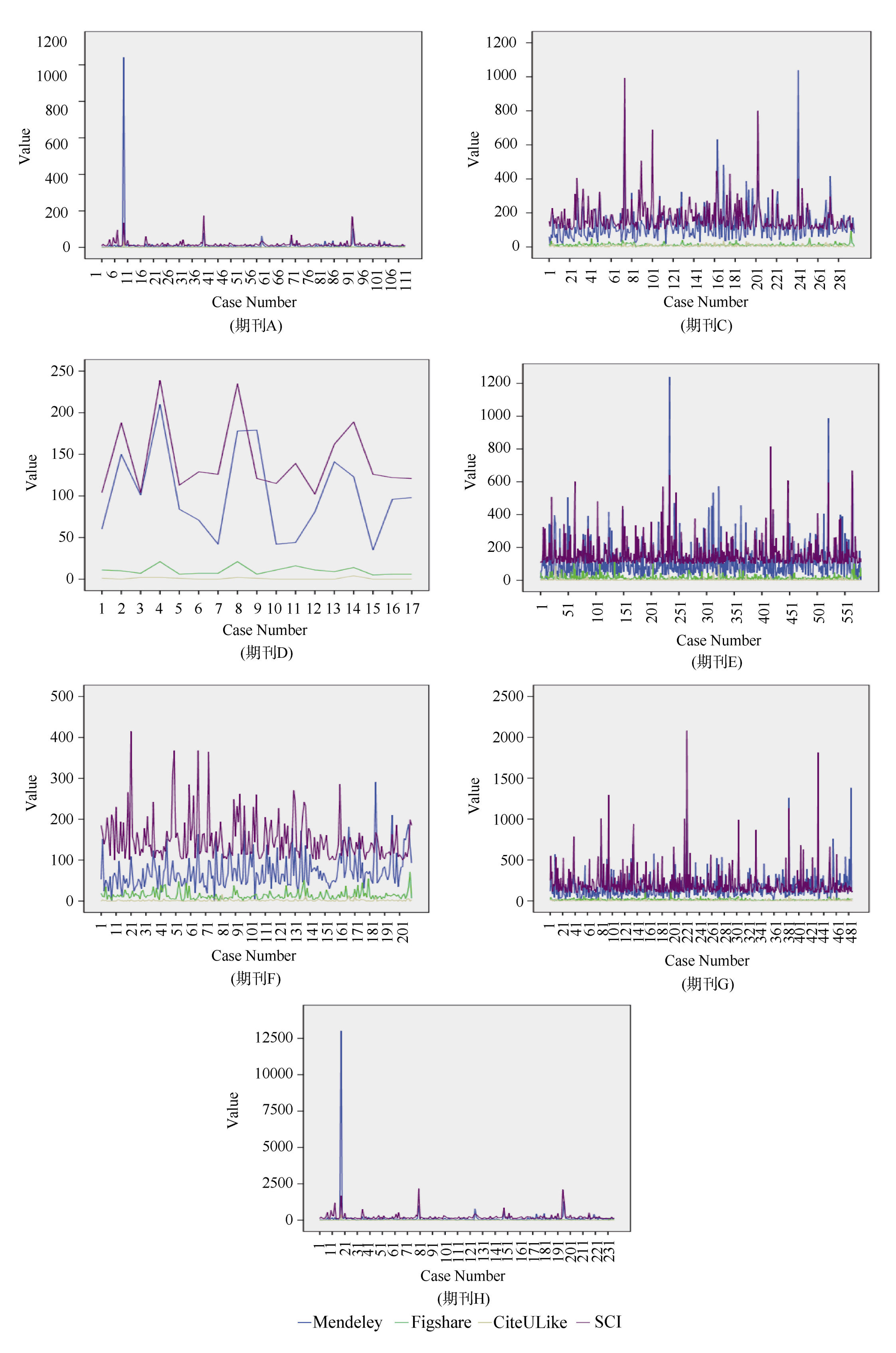


图 4 各 PLOS 期刊(A, C-H)高被引论文的在线使用与分享的变化

表 4　SCI 被引次数前 10 的论文及其在线使用与分享量

Mendeley	Figshare	CiteULike	SCI	所属期刊	Rank in SCI
994	35	6	2 170	G	1
264	37	16	2 105	E	2
1 130	19	14	2 081	F	3
140	14	11	1 814	H	4
13 015	14	389	1 689	D	5
608	26	26	1 296	C	6
1 291	36	8	1 258	E	7
168	16	2	1 191	C	8
1 260	23	60	1 132	H	9
668	23	7	1 007	G	10

表 5　Mendeley 在线使用前 10 的论文及其 SCI 引用次数

Mendeley	Figshare	CiteULike	SCI	所属期刊	Rank in Mendeley
13 015	14	389	1 689	D	1
1 382	16	29	169	E	2
1 291	36	8	1 258	E	3
1 260	23	60	1 132	H	4
1 239	14	18	641	G	5
1 150	14	106	172	E	6
1 130	19	14	2 081	F	7
1038	14	9	400	F	8
994	35	6	2 170	G	9
987	16	10	595	E	10

SCI引用次数最高的论文，根据对应的DOI，发现这些论文的主题多是关于科学研究的新发现，如关于死亡率和疾病、鼠基因组、微 RNA 靶标、人大脑皮层、DNA 条形码(如 Rank 1, 2, 4, 6, 7, 9, 10 的论文)，但是这些论文的在线使用与分享却非常低。例外的是 Rank 5(在表 5 中，Rank 1)的论文，其 SCI 引用次数居中，但是其 Mendeley 使用与 CiteULike 标注量却是最高，论文是不寻常的、令人困惑的主题，“调侃”大部分的科学发现都是错误的。

SCI 引用次数最高的论文，其在线使用与分享的量并不是最高的，同样的，以 Mendeley 使用量为例，在线使用次数最高的论文，SCI 引用次数也不是最高的。表 4 和表 5 结果显示，两项排名均在前 10 的文献有 5 篇。引用次数排名靠前的论文在 Mendeley 在线使用的量并不多，Mendeley 在线使用排名靠前的论文 SCI 被引次数也相对排名靠前。

同样的分析也适用于Figshare分享量和CiteULike标注量，结果可能略有不同。

5　研究结果与讨论

5.1　研究结果

通常认为: 0＜|r|≤0.3 为微弱相关; 0.3＜|r|≤0.5 为低度相关; 0.5＜|r|≤0.8 为显著相关; 0.8＜|r|≤1 为高度相关，只有当变量之间存在高度相关时，进行回归分析寻求其相关关系的具体形式才有意义[28]。相关性检验的结果显示: SCI 引用量和在线使用数和分享数存在正的弱相关。Mendeley 使用量与 SCI 引用之间相关性系数最大(r = 0.376, p = 0.01)，为低度相关；其次为 SCI 引用与 CiteULike 标注量的相关性系数(r = 0.252, p = 0.01)，SCI 引用与分享数(Figshare)的相关性最小(r = 0.081, p = 0.01)，为微弱相关。这说明 SCI 引用次数高的论文，其在线使用与分享的量并不高；同样的，以 Mendeley 使用量为例，在线使用次数高的论文，SCI 引用次数也不高。从 PLOS 期刊整体上看，高被引论文与其对应的在线使用与分享的相关度不高，显示出在线使用与分享的数据在揭示高影响力文献方面存在一定的不足。似乎高被引没有带来高的在线使用与分享或者是高的在线使用与分享与高被引的相关性不大。这可能与用户在线使用数据的量不足有关，应该持续关注该指标组合，待在线使用与分享的用户持续增长，能够获得充足的用户数据，继续研究其与传统引文分析评价指标的关系。有研究表明，多数高被引研究者在社会网络的存在度不高[29]，非常直接的后果可能导致社交网络平台上积累的用户数据量太小。

对高被引论文所属的期刊在线使用与分享情况进行分析的结果显示: Figshare 和 CiteULike 位于曲线底层，说明两平台上积累的用户数据量太小或者用户的活跃度太差。根据 Mendeley 峰值与 SCI 峰值的对比，部分期刊(期刊 A, C, E, H)的 Mendeley 使用量揭示了高影响力文献。此外，从使用量高的论文主题看，似乎与科学研究的新发现(高影响力文献)的主题不相关，可以看出 Mendeley 与 CiteULike 两平台除在线文献管理功能外，还兼具社交功能。

在研究高被引论文的引用次数与在线使用与分享数据相关性时，注意到 CiteULike 标注量与 Mendeley 使用量的相关性最大(r = 0.596, p = 0.01)。很大的可能

性是一篇重量级的研究文献发表后，很多科研人员或公众阅读、标注，然后才有可能进一步使用该文献，这也符合科研的一般流程。

5.2 研究不足之处

(1) 本研究的 SCI 引用数据采集自 2015 年 2 月 3 日，ALMs 数据下载的日期是 2015 年 1 月 10 日，因此，组合数据集里不包括 2 月份 ALMs 新增加的指标数据。此外 SCI 收录各个学科领域高影响力的期刊论文，其引文范围有局限性。PLOS 的这些论文的引用只是被 SCI 收录的论文的引用，如果论文未被 SCI 收录也引用了 PLOS 的论文，则本数据集中的引用数据并不包括在内。

(2) 本研究针对特定学科和特定期刊(PLOS 系列期刊)，因此，数据的全面性存在一定的欠缺。

(3) 在 CiteULike、Mendeley、Figshare 等在线学术性标签标注分享平台中，用户大多数是以英语为母语的研究者，这种偏见也与 SCI 收录的期刊多为英文期刊相一致。这样就导致这些在线使用与分享的用户数据可能单一，对结果有一定影响。

6 结 语

高被引论文并不具有高的在线使用与分享，只有部分期刊的高 Mendeley 指标揭示了高影响力文献，似乎在线的"喜欢"或者"分享"缺乏权威性和科学的可信度[30]，论文的在线使用与分享在揭示高影响力文献方面具有一定的局限性，Altmetrics 并不能替代传统的科学计量评价，只是对科研成果在社交网络上的影响力或者社会影响力的一种传播计量。传统的学术的评价(科学引用和同行评议)显然还有存在的价值，因此，构建一个多维的科研评价体系，包括 Altmetrics 和传统的文献计量，综合全面评价学术成果，势在必行。

在社交网络分享科研成果可能会提高成果的可见度和未来的引用次数。因此，科研人员抓住机会，积极参与和投身社交媒体，扩大科研成果的可见度和传播[31]，使大众更广泛地参与科研的全流程，增加对科学某些具体领域(如与人类生存息息相关的气候变化和环境科学领域)的关注。

参考文献：

[1] Brody T, Harnad S, Carr L. Earlier Web Usage Statistics as Predictors of Later Citation Impact [J]. Journal of the American Society for Information Science and Technology, 2005, 57(8): 1060-1072.

[2] Garfield E. Citation Analysis as a Tool in Journal Evaluation [J]. Science, 1972, 178(4060): 471-479.

[3] Garfield E. Citation Indexing for Studying Science [J]. Nature, 1970, 227(5259): 669-671.

[4] NISO Altmetrics Standards Project White Paper (Draft 4) [EB/OL]. (2014-06-06). [2016-06-28]. http://www.niso.org/apps/group_public/download.php/13295/niso_altmetrics_white_paper_draft_v4.pdf.

[5] Taraborelli D. Soft Peer Review: Social Software and Distributed Scientific Evaluation [EB/OL]. (2008-05-20). [2016-06-08]. http://discovery.ucl.ac.uk/8279/1/8279.pdf.

[6] Neylon C, Wu S. Article-level Metrics and the Evolution of Scientific Impact [J]. PLOS Biology, 2009, 7(11): e1000242.

[7] Priem J, Hemminger B M. Scientometrics 2.0: New Metrics of Scholarly Impact on the Social Web [EB/OL]. (2010-07-05). [2016-06-08]. http://firstmonday.org/ojs/index.php/fm/article/view/2874.

[8] Priem J, Taraborelli D, Groth P, et al. Altmetrics: A Manifesto [EB/OL]. (2010-10-26). [2016-06-08]. http://altmetrics.org/manifesto/.

[9] Galligan F, Dyas-Correia S. Altmetrics: Rethinking the Way We Measure [J]. Serials Review, 2013, 39(1): 56-61.

[10] Melero R. Altmetrics - A Complement to Conventional Metrics [J]. Biochemia Medica, 2015, 25(2): 152-160.

[11] Gordon G, Lin J, Cave R, et al. The Question of Data Integrity in Article-Level Metrics [J]. PLOS Biology. 2015, 13(8): e1002161.

[12] Ebrahimy S, Setareh F, Hosseinchari M. Assessing the Relationship Between the Alternative Metrics of Visibility and Social Bookmarking with Citation Index in PLOS Altmetrics [J]. Iranian Journal of Information Processing Management, 2016, 31(3): 845-864.

[13] Bornmann L. Alternative Metrics in Scientometrics: A Meta-analysis of Research into Three Altmetrics [J]. Scientometrics, 2015, 103(3): 1123-1144.

[14] Haustein S, Peters I, Bar-Ilan J, et al. Coverage and Adoption of Altmetrics Sources in the Bibliometric Community [J]. Scientometrics, 2014, 101(2): 1145-1163.

[15] Li X M, Thewall M, Giustini D. Validating Online Reference Managers for Scholarly Impact Measurement [J]. Scientometrics, 2011, 91(2): 461-471.

[16] Bornmann L. Usefulness of Altmetrics for Measuring the

Broader Impact of Research: A Case Study Using Data from PLOS and F1000 Prime [J]. Aslib Journal of Information Management, 2015, 67(3): 305-319.

[17] 刘春丽, 何钦成. 不同类型选择性计量指标评价论文相关性研究——基于 Mendeley, F1000 和 Google Scholar 三种学术社交网络工具[J]. 情报学报, 2013, 32(2): 206-212. (Liu Chunli, He Qincheng. Study on Correlation of Different Altmetrics Indicators for Paper Evaluation Based on Three Academic Social Networking Tools: Mendeley, F1000 and Google Scholar [J]. Journal of the China Society for Scientific and Technical Information, 2013, 32(2): 206-212.)

[18] 由庆斌, 韦博, 汤珊红. 基于补充计量学的论文影响力评价模型构建[J]. 图书情报工作, 2014, 58(22): 5-11. (You Qingbin, Wei Bo, Tang Shanhong. Evaluation Model Construction to Evaluate Article's Influence Based on Altmetrics [J]. Library and Information Service, 2014, 58(22): 5-11.)

[19] 宋丽萍, 王建芳, 王树义. 科学评价视角下 F1000、Mendeley 与传统文献计量指标的比较[J]. 中国图书馆学报, 2014, 40(4): 48-54. (Song Liping, Wang Jianfang, Wang Shuyi. Observation of F1000, Mendeley and Traditional Bibliometric Indicators from the Perspective of Scientific Assessment [J]. Journal of Library Science in China, 2014, 40(4): 48-54.)

[20] 刘晓娟, 周建华, 尤斌. 基于 Mendeley 与 WoS 的选择性计量指标与传统科学计量指标相关性研究[J]. 图书情报工作, 2015, 59(3): 112-118. (Liu Xiaojuan, Zhou Jianhua, You Bin. Study on Correlation Between Altmetrics Indicators and Traditional Scientometric Indicators Based on Mendeley and WoS [J]. Library and Information Service, 2015, 59(3): 112-118.)

[21] 金玮, 赵蓉英, 殷鸽. 用户在社会化引文软件中的阅读数据积累程度与有效性分析——以 Altmetrics 指标为例[J]. 现代图书情报技术, 2015(11): 75-81. (Jin Wei, Zhao Rongying, Yin Ge. An Analysis of the Accumulation State and the Validity of User Readership Data in Online Reference Managers——Take the Indicators of Altmetrics as an Example [J]. New Technology of Library and Information Service, 2015(11): 75-81.)

[22] Mohammadi E, Thelwall M. Mendeley Readership Altmetrics for the Social Sciences and Humanities: Research Evaluation and Knowledge Flows [J]. Journal of the Association for Information Science and Technology, 2014, 65(8): 1627-1638.

[23] Shrivastava R, Mahajan P. Relationship Between Citation Counts and Mendeley Readership Metrics: A Case of Top 100 Cited Papers in Physics [J]. New Library World, 2016, 117(3-4): 229-238.

[24] Hane P J. Sharing Research Data—New Figshare for Institutions [EB/OL]. (2013-09-16). [2016-06-28]. http://www.against-the-grain.com/2013/09/sharing-research-data-new-figshare-for-institutions-2/.

[25] Hahnel M. Figshare: A New Way to Publish Scientific Research Data [EB/OL]. (2012-01-18). [2016-06-08]. http://blog.wellcome.ac.uk/2012/01/18/figshare/.

[26] Cumulative PLOS ALM Report-January 2015 [DS/OL]. [2015-02-01]. http://dx.doi.org/10.6084/m9.figshare.1367546.

[27] 匡登辉. 高被引论文的论文级计量分析[J]. 情报杂志, 2016, 35(5): 149-154, 194. (Kuang Denghui. ALMs Analysis of the Most Highly Cited Articles [J]. Journal of Intelligence, 2016, 35(5): 149-154, 194.)

[28] 卢冶飞, 孙忠宝. 应用统计学[M]. 第 2 版. 北京: 清华大学出版社, 2015: 36-48. (Lu Yefei, Sun Baozhong. Applied Statistics [M]. The 2nd Edition. Beijing: Tsinghua University Press, 2015: 36-48.)

[29] Masbleda A, Thelwall M, Kousha K, et al. Do Highly Cited Researchers Successfully Use the Social Web? [J]. Scientometrics, 2014, 101(1): 337-356.

[30] Cheung M K. Altmetrics: Too Soon for Use in Assessment [J]. Nature, 2013, 494(7436): 176.

[31] Bar-Ilan J, Haustein S, Peters I, et al. Beyond Citations: Scholars' Visibility on the Social Web [EB/OL]. [2016-06-28]. http://arxiv.org/ftp/arxiv/papers/1205/1205.5611.pdf.

利益冲突声明:

作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail: dhkuang@nankai.edu.cn。

[1] 匡登辉. alm_stats_2015_01_10.csv. PLOS ALMs.

[2] 匡登辉. HCP.xlsx. PLOS 高被引论文(WoS)数据.

[3] 匡登辉. Combinatorial.xls. 高被引论文(WoS)与 ALMs 在线使用与分享组合数据.

[4] 匡登辉. PLOS Serials.xlsx. PLOS 系列期刊论文在线使用与分享数据.

收稿日期: 2016-07-08
收修改稿日期: 2016-09-08

Analyzing Online Usage & Sharing of Highly Cited Papers

Kuang Denghui
(Nankai University Library, Tianjin 300071, China)

Abstract: [**Objective**] This empirical case study aims to validate the effectiveness of using Altmetrics indicators to identify high quality articles. [**Methods**] First, we retrieved the online usage and sharing data of highly cited papers published by the PLOS journals from social platforms (i.e., CiteULike, Mendeley and Figshare). Second, we examined relationship between these Altmetrics and SCI citation counts of the target papers. [**Results**] The correlation coefficient between the SCI citation data and the Altmetrics generated by Mendeley was strong ($r = 0.376$, $p = 0.01$). Meanwhile, the other two correlation coefficients were weaker. The online usage data from Mendeley might help us identify high impact literature published by specific journals. [**Limitations**] This research only investigated a few subjects covered by the PLOS serial journals. More research is needed to check the relationship between Altmetrics and citation counts in other fields. **[Conclusions**] Online usages & sharing data from CiteULike, Mendeley and Figshare might not be able to effectively identify the high impact literature.

Keywords: Highly cited papers　Online usages　Mendeley　CiteULike　Sharing　Figshare

新技术促使澳大利亚的图书馆成为社区中心

学界一些新的研究表明，地方图书馆正面临着不断变化的社区需求。由 Civica 公司和悉尼科技大学公共政策与治理研究所进行的题为“图书馆作为公共空间的价值”的深入调查发现：未来的图书馆将成为社区支持中心。在许多情况下，图书馆需要 7×24 小时开放。

题为“图书馆作为公共空间的内在价值”的报告在 Civica 博览会上由悉尼科技大学公共政策与治理研究所主任 Roberta Ryan 副教授发起。

该调查统计了澳大利亚和新西兰的图书馆的管理者、老员工，以及地方政府高级管理人员的意见。受访者一致认为侧重于沉默阅读和学习的图书馆已经过时。

Civica 公司图书馆与教育解决方案总监 Simon Jones 说：“这项研究表明，图书馆不仅需要改变，更应该不断寻找新的方法和新的模式，以更好地服务于社会需求。”

该调查指出，激进的新方法，如在图书馆提供企业孵化器设施，将有助于社区新企业的诞生。只要提供免费 WiFi 和计算机服务，图书馆就可以给潜在的企业家提供一个他们需要的基础条件，而这是他们在公开市场上无法负担的。

“7×24 小时开放的图书馆在许多情况下是无人值守的，这已经在包括瑞典在内的几个国家开始运作，这种模式很可能很快就会在澳大利亚的一些城市进行试用。” Simon Jones 指出。

图书馆的未来将是一站式的，为社区提供失业援助、健康咨询、社区学习和商业发展等支持。

虽然图书馆员已经接受了数字时代的挑战，但他们没有离开他们的核心责任，即，提供阅读图书和学习的机会。 然而，他们都同意，在数字时代，有许多新方法能为社区中的更多人提供资源。他们认为，提供阅读图书和学习机会的核心功能将保持不变，但一系列新服务的开展仍需围绕该核心功能。

超过 77%的受访者表示，图书馆使用免费 WiFi 为读者提供一般互联网使用是现代图书馆的重要功能。

受访者还认为，图书馆如果提供更多基于社区的空间将有助于促进社会凝聚力和社区交互，防止社会孤立。

(编译自: https://librarytechnology.org/news/pr.pl?id=21985)

(本刊讯)